DOCUMENT RESUME

ED 297 071 TM 011 989

AUTHOR Kieren, Dianne K.; Munro, Brenda
TITLE The Observational Recording Dilemma.
SPONS AGENCY Social Sciences and Humanities Research Council of Canada, Ottawa (Ontario).
PUB DATE [85]
GRANT 410-84-0256
NOTE 26p.; For related documents, see TM 011 987-988.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Comparative Analysis; *Decision Making; *Family Life; Family Problems; *Interrater Reliability; *Observation; Problem Solving; *Research Methodology; Videotape Recordings
IDENTIFIERS Coding; *Observation Techniques; *Transcription Method

ABSTRACT

Decision making about an observational recording system for family interaction research is crucial. Alternative coding-recording methods and combinations thereof are discussed, including: (1) paper-and-pencil on-site method; (2) video-tapes; (3) paper-and-pencil and mechanical coding devices; (4) transcripts; and (5) transcripts combined with video-tapes. Data from a family problem solving study illustrate the type of questions that must be answered to arrive at a decision for a research study. Time requirements, number of units coded, interrater reliabilities for transcript methods, and data profile comparisons are reported. The on-site method was the least costly, but revealed the most discrepant data profile compared with the other methods. Transcript methods were most costly, but also provided higher levels of precision in detailing family interactions. Video alone appeared to offer the best choice for coding-recording considering cost, comprehensiveness, and accuracy. Further testing of such methods is suggested to assist researchers in making decisions about coding-recording methods. Five tables are included. (TJH)

THE OBSERVATIONAL RECORDING DILEMMA

Dianne K. Kieren, Ph.D.
Professor
and
Brenda Munro, Ph.D.
Assistant Professor
and McTaggart Fellow

Department of Family Studies
801 General Services
Faculty of Home Economics
University of Alberta
Edmonton, Alberta, Canada
T6G 2E2

Research was supported by an SSHRC grant 410-84-0256. Thanks are expressed to Susan Horner for her assistance in data analysis.

AUTHORS

Dianne K. Kieren, Ph.D. Professor of Family Studies,
Family Studies Department,
Faculty of Home Economics,
University of Alberta,
801 General Services Building,
Edmonton, Alberta, Canada,
T6G 2E2

Brenda Munro, Ph.D. Assistant Professor of Family Studies,
and McTaggart Fellow,
Family Studies Department,
Faculty of Home Economics,
University of Alberta,
801 General Services Building,
Edmonton, Alberta, Canada,
T6G 2E2

ACKNOWLEDGMENTS

This research was supported by a Social Sciences and Humanities Council of Canada Grant (410-84-0256) entitled "Family Problem Solving in Families with and without an Adolescent Diabetic". Thanks are expressed to Susan Horner for her assistance in data analysis.

ABSTRACT

THE OBSERVATIONAL RECORDING DILEMMA

Decision making about an observational recording system for family interaction research is crucial. Issues of cost and quality of data are primary. The present paper discusses four alternative coding-recording methods: on-site: paper and pencil; video only: paper and pencil and mechanical coding devices; transcript only: paper and pencil; transcript plus video: individual coder, paper and pencil and consensus. Data from a family problem solving study are used to illustrate the type of questions which must be answered to arrive at a decision for a research study. Time requirements, number of units coded, interrater reliabilities for transcript methods, and data profile comparisons are reported. The on-site method was the least costly but revealed the most discrepant data profile compared with the other methods. Transcript methods were most costly but also provided higher levels of precision in detailing family interaction. Video alone appeared to offer the best choice for coding-recording considering both cost, comprehensiveness and accuracy. Further testing of such methods is suggested to assist researchers in making decisions about coding-recording methods.

THE OBSERVATIONAL RECORDING DILEMMA

Introduction

Recently, the amount of observational study of family and marital functioning has increased (See Filsinger, 1983; Steinglass, 1979; Steinglass & Tislenko, 1983). The growing desire to gain the most accurate picture of family interaction combined with the decline in available grant funds necessitates careful selection of research design, methodology and type of analysis. Researchers must make many decisions in order that the ensuing study might be the most representative possible within economic constraints. Thus, doing observational research involves the resolution of many methodological issues.

One issue involves deciding upon the most appropriate, accurate and manageable recording system given the nature of the interaction being measured, the type and complexity of the code and the funding resources available. A number of choices are available: on-site observation and coding using either a paper and pencil or a mechanical recording device; observation and coding from videotape with or without transcription; coding from audiotape with or without transcription; or consensus coding using video, audio and/or transcript. Each method has its advantages and limitations. The use of transcripts increases the time and other costs. This method, however, provides greater precision in the coding process.

The newer mechanical coding instruments (e.g. OS3 and Data Myte) may be used on-site and/or with video and have the

potential of eliminating the transcribing step.

These methods require different training processes and challenge the researcher in assessing coding reliability especially when the code is a complex one. The question posed in this paper is: "Do these less costly coding-recording methods obtain the same picture of family interaction as more costly ones?"

Frequently the evaluation of different methods has been limited to a focus on intercoder consistency or reliability. Hartmann and Gardner (1981) defined the term observer reliability in terms of two related yet distinct paradigms, observer accuracy and interobserver agreement, reliability or consistency. The accuracy paradigm compares what is assumed to be a flawed source of data with what is assumed to be an accurate or error less evaluation of an occurrence. Interobserver agreement compares two data sources that are assumed to be flawed with each other. In both cases, one code of a particular behavioral unit is compared with another code of the same behavioral unit. However, this type of comparison is not always possible when observing ongoing family interaction without the help of a transcript. When coders must decide upon the coding unit as well as the type of behavior, the number of interactions coded may vary and interrater reliabilities may be flawed.

Within the present investigation data profile reliability will be used in addition to assessing coder accuracy and agreement, in the comparison of recording methods. That is,

while the number of coded interactions might differ among recording techniques, the question posed, is whether the picture or profile of the family's interaction is similar. If there is data profile reliability, the proportion of codes in different categories should be the same or similar regardless of the method used. For example, a researcher using a transcript of an observational session is usually asked to code already identified coding units while an on site coder must identify coding units from the flow of interaction. Thus the coder might identify only a sampling of the coding units identified on a transcript. In addition, the on-site coder has no chance for a re-review. Even so, the portrait or profile obtained of family interaction utilizing these two methods could be similar and would provide a compelling rationale for the choice of this recording method over others.

Thus, the purpose of this investigation was to conduct a limited test of these recording methods in an ongoing study of family problem solving interaction to provide a better rationale for the selection and use of a particular recording method in subsequent studies. Fully trained coders were used to determine the relative accuracy and time requirements of various recording methods. Comparisons of the methods were made based on the following questions:

1) How do the number of ratings vary using different coding-recording methods?

2) How does the time investment for each method compare?

3) How does the interrater reliability of the different recording methods using transcripts compare to a consensus code?

4) How do the family profiles of family problem solving interaction which emerge from these methods differ?

A Review of Recording Methods

Perhaps the most frequently used method of recording is the on-site paper and pencil technique. Coders using only paper and pencil are faced with the formidable task of being able to identify and code behaviors in many response categories in a split second of time. Because the difficulty of this task has to be noted, commercial companies have come to the aid of researchers (Holm, 1981) by producing machines which theoretically allow the coder to more rapidly and accurately, record observed behavior. Perhaps the most frequently used machines are the Datamyte, the OS3 and the SSR System 7 (Filsinger, 1983).

Another observational aid is the video recorder. Utilization of video recorded family interaction, allows coders to utilize many recording techniques. For example, coders might code using paper and pencil, machines, or transcripts of the session. The distinct advantage of video is being able to turn the tape back in order to verify findings.

While all of these recording techniques are viable, each has its distinct strengths and weaknesses which must be weighed by individual researchers dependent on their particular needs.

On one hand, researchers want to gather information that is as reliable as is humanly possible. This means finding a technique where an optimal amount of family interaction information can be gleaned from observational settings with a methodological provision for testing the reliability of obtained information.

Hartmann and Gardner (1981) discuss the importance of reliability in research in order to maintain as they state, the "quality of our research" (p. 184). It is important that measurement be 1) accurate - correspond to the observed events, 2) precise - that random error noise be reduced, and 3) generalizable - that obtained observations are representative of the real world and thus could be duplicated by other researchers.

However, realistically, the economy of research projects is also an issue. Observational research is time consuming, both prior to and subsequent to the family interaction sessions. In most cases time and money constraints are a reality and researchers must do the best research they can within financial and time boundaries. Thus research needs to be done without sacrificing the accuracy, precision and generalizability of the information obtained.

Methodology

Coding Methods

The present paper includes a report of the analysis of two problem situations comprising approximately 200 behavioral

codes in each situation from a larger study of family problem solving interaction of three person family groups (Kieren & Hurlbut, 1985). In this analysis four different methods of coding and recording family problem solving interaction were compared in order to answer the research questions posed. The methods included:

1) coding on-site: one coder using a data logging instrument and one using a paper and pencil technique;
2) coding using video: one coder using the data logging instrument and the other using paper and pencil; this was done one week after the initial on-site coding session;
3) transcript only: paper and pencil;
4) coding using transcript plus video: both coders using paper and pencil recording methods; and both coders using paper and pencil techniques and discussing until consensus was reached.

Coders were randomly assigned to the different methods in procedures one and two.

Problem Solving Interaction

The problem solving task for the family involved a revealed difference task developed by Kieren, Hurlbut, Gora, & Lehman (1985). Each family member was asked to independently answer a series of questions about nine common family problem solving situations. The questions were: What is the problem? Has a situation like this one ever happened in your family? Who is most responsible for the problem? What should be done? Who

should make the final decision? Once each person in the family had independently answered the questionnaire their responses were reviewed by the researchers. Three of these situations were selected based upon salience for the family and some level of revealed difference between member's responses. The family was allowed up to ten minutes to discuss each situation and come to a family consensus on answers to the questions. The interaction was videotaped. The present comparison was based upon the analysis of two of these situations since the multiple coding of these situations included the coding of over 1600 behavioral units.

General Coding Procedures

A twenty six category code was developed which specifically tapped family problem solving behavior (Kieren, 1985). Coders were required to code each complete thought unit using a four digit code. The code represented 1) the sender, 2) the receiver and 3-4) the specific two digit code for each problem solving behavior.

Two trained coders were utilized to code the data. Training involved an extensive process of an average of 55 hours. This involved studying the code, achieving sufficiently high content accuracy (90%), practicing using transcripts as well as transcripts plus video, learning to use a data logging instr [illegible] (OS3), practicing using a data logging instrument and [illegible] with trainers. Interrater reliability was dete [illegible] during training using a kappa statistic (Cohen,

1960) applied to a trial when coders used video plus transcripts since such a method requires comparison of similar coding units. The reliability achieved during training was .75 which compares favorably with levels reported by other interaction researchers using a complex molecular code (Raush, Barry, Hertel & Swain, 1974).

Comparative Coding Procedures

After training was completed, each coder was randomly assigned to code the individual family problem solving situations. Both coders were on-site to code the problem solving interactions; one used the OS3 instrument and the other the paper and pencil technique. The family interaction for each problem vignette was also videotaped. A different research assistant transcribed the two problem solving interactions during the next week following the on-site coding. Problem solving situation #1 involved 6 minutes 40 seconds of interaction whereas problem solving interaction #2 constituted 5 minutes 20 seconds. During the next two weeks, the remaining coding and recording methods were completed. Some bias is built into this method in that the coders were exposed to the same family interaction several times and thus may have improved in their level of reliability, however using additional coders would have required costly training. Thi ..ay have been minimized by the fact they were also coding other families during this same period.

Results

Coding was completed as planned with one exception. During the on-site coding an editing problem developed with the use of the OS3 which made it necessary to drop this technique from the on-site analysis.

Number of Ratings

All techniques using transcripts had a defined number of coding units, 113 in situation 1, and 108 in situation 2.

Place Table 1 about here

For the remaining recording methods, the fewest number of units were coded using the on-site paper and pencil recording method. The number of units coded in situation #1 was 37% of the identified units on the transcript. For situation #2, the percentage was 47%. The recorders identified slightly more coding units using video than they did in direct observation. Using the paper and pencil recording method with video the recorders obtained 78% and 87% of the units identified in the transcript for situations 1 and 2. Using the OS3 recording device with video the percentages were 70% and 66% respectively.

Interrater Reliability

Cohen's (1960) coefficient of agreement was calculated for

each recording method using a transcript. Each coder's rating was compared to the ratings arrived at by consensus. Table 2 reports the data. This method is a conservative estimate of reliability compared with simple percentage agreement. Reliabilities were very high and may reflect some inflation due to the repeated viewing of transcripts or videos.

Place Table 2 about here

Time Requirements

Table 3 summarizes the time estimates arrived at for the differing methods. On-site coding was the most economical, time-wise. Use of video without transcript and a paper and pencil recording method involved a minimum of ten minutes to set up the video and at least three passes of the tape for coding. Using a mechanical recording device (OS3) had an additional time requirement to edit where necessary and to dump the data to a storage tape (30 minutes).

Methods requiring a transcript required much more time. Transcribing is time consuming. We have estimated it takes a minimum of 2 1/2 hours to produce a clean, accurate copy of up to 10 minutes of tape. In addition it takes 30 minutes to unitize the transcript. An experienced coder can code up to 115 items in approximately 30 minutes. Using a transcript as well as video adds the possibility of re-reviewing the tape while coding. We estimate this takes about 4 to 5 times the time estimate for video alone.

Consensus coding demands that coder's continue discussing an item until a similar coding category is agreed upon. Based on a disagreement rate of approximately 15% and discussion time of 3 minutes per coding unit, an additional one hour of coding time was added beyond that reported for transcript and video. For researchers, such time requires additional funding.

Profile Comparisions

Data comprising the entire 26 category problem solving code were summarized into seven summary categories: S_1: Fragments; S_2: Identification of Problem; S_3: Alternative Generation, S_4: Evaluation; S_5: Resolution Mechanisms; S_6: Proposed Decisions; S_7: Meta-problem solving. Relative percentages based upon total units in the summary code compared with the total units coded were calculated for each summary code. This represented the family profile used to compare the recording methods (Tables 4 & 5).

The methods using transcripts revealed profile results which were remarkably similar: The video data had some noteable discrepancies when compared with the transcript codes particularly for summary codes S_1, S_3, S_5, and S_6 but in general the differences were not great. The on-site coding was most discrepant with an overuse of certain codes (note the high percentage reported for S_2 compared with the other methods).

Discussion

The selection of a coding/recording method is a crucial

decision for a researcher studying family interaction. Even an arm chair comparison of methods reveals that less stringent recording methods have advantages. It would be expected that they would be less costly in terms of time and therefore money. Reduced time and money make the cost per subject less formidible and allows a larger sampling of the population. This would be more appealing to funding agencies. The question which remains unanswered by such a cursary analysis is whether the researcher sacrifices quality data for the cost savings. The present paper provided an example of a systematic analysis which could be used to compare coding recording methods. The data profile of family problem solving interaction which resulted from the use of different recording methods was compared in addition to several other variables. The profile results indicated that on-site observation and coding was the least desirable method. Not only were significantly fewer codes obtained but the data profile revealed was very different from that obtained by other methods. Training did not appear to ameliorate this discrepancy. Both of the on-site coders had been involved in an extensive training process which involved training to identify the coding unit as well as to apply the code.

This data contrasts with that reported by Bench, Hoffman and Wilson (1974) and Bench and Wilson (1975, 1976) comparing live versus video recording on neonates, 6 week-old and 6 month-old infants responding to sound stimulation. They report that little or no information is lost using live observation

and in fact, observation of 6 month-olds may even be somewhat inferior using video. Sackett, Ruppenthal and Gluck (1978) reported that live recording catches 90% of information extracted from repeated viewing of monkeys interacting on a 16mm film.

Several factors could account for these differences. One is coder anxiety. On-site coding has an added element of coder anxiety in that this is a once and for all chance to collect the data. In this study the coder had to make a decision regarding the unit to be coded as well as assigning one of 26 possible problem solving codes to the unit. This involves split second decision making. The task may be more complex than in previously reported studies which observed more molor behavior. Another more important factor is the fact that 3 person family groups were being observed by one observer. When three persons are interacting interruptions and overlaps of communication naturally occur. The on-site recording and coding difficulties of poor sound transmission to the coder cannot be compensated for by a re-review. The factors may decrease both comprehensiveness (number of units coded) as well as accuracy (data profile).

Use of the video increased the level of comprehensiveness as well as accuracy, even without a transcript. There were, however, some sacrifices in the use of the technique. It would appear that fragments (incomplete statements) were identified less frequently than in techniques using a transcript. There were also some small differences in reported proportions.

All of the methods using transcripts were most similar in comprehensiveness and accuracy. These were also the most costly.

The systematic comparison of recording techniques reported in this paper provided support for the use of video without transcription for the present study of 3 person family problem solving interaction. Because the data were derived from a limited sample of approximately 100 behavioral units in each of two family problem solving interactions further comparisons are warranted before generalizing to other interactional research.

References

Bench, J., Hoffman, E., & Wilson, I. (1974). A comparison of live and videorecord viewing of infant behavior under sound stimulation I: Neonates. Developmental Psychobiology, 7, 455-464.

Bench, J., & Wilson, I. (1975). A comparison of videorecord viewing of infant behavior under sound stimulation II: Six week-old infants. Developmental Psychobiology, 8, 347-353.

Bench, J., & Wilson, I. (1976). A comparison of live and videorecord viewing of infant behavior under sound stimulation III: Six month-old babies. Developmental Psychobiology, 9, 297-303.

Cohen, J. (1960). A coefficient of agreement for nominal scales. Educational & Psychological Measurement, 20(1), 37-40.

Filsinger, E. E. (1983) Marriage and family assessment. Beverly Hills, CA: Sage Publications.

Filsinger, E. E. (1983). A machine-aided marital observation technique: The dyadic interaction scoring code. Journal of Marriage and the Family, 45(3), 623-632.

Hartman, D. P., & Gardner, W. (1981). Considerations in assessing the reliability of observation. In E. Filsinger & R. A. Lewis (Eds.), Assessing marriage. Beverly Hills, CA: Sage Publishers.

Holm, R. (1981). Using data logging equipment. In E. Filsinger & R. Lewis (Eds.), Assessing marriage. Beverly Hills, CA: Sage Publishers.

Kieren, D. (1985). Problem solving behavioral inventory code. Edmonton, Department of Family Studies, (mimeographed).

Kieren, D., & Hurlbut, N. (1985). Final report: Problem solving interaction in families with and without a diabetic adolescent. Edmonton, Department of Family Studies (mimeographed).

Kieren, D., Hurlbut, N., Lehman, D., & Gora, R. (1985). Family revealed difference instrument (RDI). Edmonton, Department of Family Studies (mimeographed).

Raush, H., Barry, W., Hertel, R., & Swain, M. (1974). Communication conflict and marriage. San Francisco: Jossey-Bass Publishers.

Sackett, G. P., Ruppenthal, G. C., & Gluck, J. (1978). Introduction: An overview of methodological and statistical problems in observational research. In G. P. Sackett (Ed.), Observing Behavior. Baltimore: University Park Press.

Steinglass, P. (1979). The home observation assessment method (HOAM): Real-time naturalistic observation of families in their homes. Family Process, 18, 337-354.

Steinglass, P., & Tislenko, L. (1981). The home observation assessment method. In E. Filsinger & R. Lewis (Eds.), Assessing marriage. Beverly Hills, CA: Sage Publishers.

TABLE 1

Number of Ratings Compared Over Recording Methods

Technique	Number of Ratings	
	Situation #1	Situation #2
On-site		
OS3 (dropped)	--	--
Paper and Pencil (Coder #2)	42	51
Video		
OS3	79	71
Paper and Pencil (Coder #1)	88	94
Transcript Only		
Coder #1	113	108
Coder #2	113	108
Transcript and Video		
Paper and Pencil (Coder #1)	113	108
Paper and Pencil (Coder #2)	113	108
Transcript and Video		
Consensus - (Both Coders)	113	108
	---	--- *

*Consensus coding is done jointly.

TABLE 2

Interrater Reliability Estimates: Transcript Methods

Coding Method	Coder	K Situation #1	K Situation #2
Transcript	1	0.957	.905
	2	.9148	.911
Video & Transcript	1	.9680	.931
	2	.8849	.8074

TABLE 3

Time[1] Required by Coding Method

	Estimated Time	
	Situation #1	Situation #2
On-site OS3 Paper and Pencil	6:40 6:40	5:20 5:20
Video only* OS3 Paper and Pencil *3 passes, 10 min set up, dumping for OS3	50:00 30:00	46:00 26:00
Transcript only* Paper and Pencil Coder 1 Paper and Pencil Coder 2 *Includes transcribing, cleaning, unitizing, coding	240:00 210:00	210:00 180:00
Transcript and Video* Paper and Pencil Coder 1 Paper and Pencil Coder 2	290:00 270:00	270:00 260:00
Transcript and Video* Consensus Code *Includes transcribing, cleaning, unitizing, coding, re-reviewing, and discussing	360:00	345:00

[1]Time is recorded in minutes and seconds

TABLE 4

Relative Percentages of Codes by Summary Code by Different Recording Methods

Situation #1

Coding Method	Coder	S_1	S_2	S_3	S_4	S_5	S_6	S_7
On Site								
OS3 (dropped)		--	--	--	--	--	--	--
Paper and Pencil	#2	0	68	0	22	10	0	0
Video								
OS3		13	38	11	23	10	5	0
Paper and Pencil	#1	13	38	8	15	15	11	0
Transcript Only								
Paper and Pencil	#1	22	38	4	17	9	10	0
Paper and Pencil	#2	24	37	6	17	8	8	0
Transcript and Video								
Paper and Pencil	#1	21	36	4	17			
Paper and Pencil	#2	22	38	6	15	8	10	0
Transcript and Video Consensus		22	40	4	16	8	10	0

TABLE 5

Relative Percentages of Codes by Summary Code
by Different Recording Methods

Situation #2

Coding Method	Coder	S_1	S_2	S_3	S_4	S_5	S_6	S_7
On Site								
OS3								
Paper and Pencil	#2	2	67	2	22	8	0	0
Video								
OS3		13	42	11	24	4	6	0
Paper and Pencil	#1	11	39	5	20	11	14	0
Transcript Only								
Paper and Pencil	#1	15	39	8	22	5	11	0
Paper and Pencil	#2	13	44	5	22	5	11	0
Transcript and Video								
Paper and Pencil	#1	11	43	3	22	7	14	0
Paper and Pencil	#2	13	40	8	22	6	11	0
Transcript and Video Consensus		13	41	6	23	5	13	0